



## Coping with Abundant Missing Entries in Phylogenetic Inference Using Parsimony

Mark Wilkinson

*Systematic Biology*, Volume 44, Issue 4 (Dec., 1995), 501-514.

Stable URL:

<http://links.jstor.org/sici?sici=1063-5157%28199512%2944%3A4%3C501%3ACWAMEI%3E2.0.CO%3B2-8>

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://uk.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*Systematic Biology* is published by Society of Systematic Biologists. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://uk.jstor.org/journals/ssbiol.html>.

---

*Systematic Biology*

©1995 Society of Systematic Biologists

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact [jstor@mimas.ac.uk](mailto:jstor@mimas.ac.uk).

©2002 JSTOR

## COPING WITH ABUNDANT MISSING ENTRIES IN PHYLOGENETIC INFERENCE USING PARSIMONY

MARK WILKINSON

*School of Biological Sciences, University of Bristol, Bristol BS8 1UG, England<sup>1</sup>*

**Abstract.**—When cladistic data sets include taxa with abundant missing entries, parsimony analysis may yield multiple equally optimal trees and necessitate the use of consensus methods to summarize relationships that are common to the multiple trees. Determination of those relationships that are common to the equally parsimonious trees and are thus unambiguously supported by the parsimonious interpretation of the data may not be possible using consensus methods that are widely employed by systematists. Thus, missing data may have an obfuscatory effect upon phylogenetic relationships. This problem can be ameliorated or overcome by adopting a strategy of safe taxonomic reduction. In this approach, taxa that can have no effect upon the relationships inferred for other taxa but that may increase the numbers of equally most-parsimonious trees are identified. Eliminating such taxa through the application of a series of safe deletion rules may reduce the number of equally most-parsimonious trees and thereby facilitate the consensus representation of unambiguous relationships supported by the data. The methods are illustrated by reanalysis of cladistic data for the Saurischia. [Phylogeny; parsimony; underdetermination; equivalence; safe taxonomic reduction; consensus; Saurischia.]

Missing entries are a common feature of cladistic data sets. Characters commonly need to be coded as missing for some taxa because the relevant parts are not preserved or have not been examined. In addition, some characters may be inapplicable in some taxa (e.g., tooth characters in birds and turtles, limb characters in snakes and caecilians) and must also be coded as missing for these taxa (Platnick et al., 1991). In both cases, the coding reflects ignorance about the relationships of a character state (known or unknown) to other character states (Wilkinson, 1992a).

Amounts of missing data vary greatly from study to study. Although missing entries are not rare in neontological cladistic data, the greatest abundance of missing entries are encountered in paleontological data sets that include poorly preserved fossils as terminal taxa. In a survey of 30 paleontological and 51 neontological data sets drawn from the literature on tetrapod phylogeny, the average level of missing data, expressed as a percentage of the total number of data points, is 12.6% for the paleontological data (range, 0–52%) and 2.29% (0–12.25%) for the neontological data.

When analyzed using parsimony, missing entries contribute nothing positive to the analysis. According to Swofford (1990: 17), “only those characters that have non-missing values will affect the location of any taxon on the tree” In turn, the placement of taxa on the tree, as determined by informative characters, can lead to parsimonious reconstructions of the values of missing entries. The principles of parsimonious reconstruction of missing entries underpin the approach to the reconstruction of unpreserved fossil morphology recently advocated by Bryant and Russell (1992).

However, missing entries are not always innocuous to phylogenetic inference using parsimony analysis. It has been widely recognized that including poorly known taxa (with much missing data) in computerized parsimony analyses may lead to a dramatic increase in the numbers of equally most-parsimonious trees (MPTs) and an accompanying loss of resolution in consensus trees that are used to summarize common elements of the MPTs (Gauthier, 1986; Nixon and Wheeler, 1992; Novacek, 1992a, 1992b; Wilkinson, 1992a; Wilkinson and Benton, 1995). Several authors have therefore excluded poorly known taxa from parsimony analyses to minimize the num-

<sup>1</sup> E-mail: mark.wilkinson@bris.ac.uk.

ber of MPTs and maximize the resolution of relationships.

Here, I describe problems that may arise from an abundance of missing entries in cladistic data sets used in parsimony analysis and develop a simple strategy, termed *safe taxonomic reduction*, for ameliorating them. In essence, safe taxonomic reduction involves the a priori identification and removal of those terminal taxa that, by virtue of their abundance of missing entries, increase the number of MPTs but that can be removed from an analysis without any danger of affecting relationships among the remaining terminal taxa that are supported by the parsimonious interpretation of the complete data. The approach is illustrated through a reanalysis of Gauthier's (1986) data for the Saurischia.

#### THE PROBLEM

The basic problem to be addressed has two parts: (1) data sets with abundant missing entries for some taxa often support numerous MPTs, and (2) currently available consensus methods have limited success in representing relationships that are common to multiple MPTs. By increasing the number of MPTs and necessitating the use of consensus methods, missing entries may obscure relationships among taxa that do not have abundant missing entries. My aim is to develop methods that can be employed in conjunction with available computerized parsimony packages to make these obscured relationships recoverable from the data.

The problem can best be explained by reference to a simple example. The hypothetical data set in Table 1 includes six taxa (A–F) that are scored for nine characters and a seventh taxon (X) that is poorly known and is coded with missing entries for the majority of the characters. Considering only the completely known taxa (A–F), the data support a single, fully resolved MPT, with no homoplasy in any of the characters (Fig. 1a). Character 3 provides unambiguous support for the hypothesis that taxon X is more closely related to the C–F clade than to either A or B. Similarly, character 7 provides unambiguous sup-

TABLE 1. Hypothetical character data for six well-known taxa (A–F) and one poorly known taxon (X) supporting the trees in Figure 1, together with parsimonious reconstructions (1–7) of the missing character states of taxon X corresponding to its possible positions in Figure 1a.

	Characters <sup>a</sup>								
	1	2	3	4	5	6	7	8	9
Taxon									
A	0	0	0	0	0	0	0	0	0
B	1	0	0	0	0	0	0	0	0
C	1	1	1	0	0	0	0	1	0
D	1	1	1	1	1	0	0	0	1
E	1	1	1	1	1	1	1	0	0
F	1	1	1	1	1	1	1	0	0
X	?	?	1	?	?	?	0	?	?
Reconstruction for X									
1	1	0	1	0	0	0	0	0	0
2	1	1	1	0	0	0	0	0	0
3	1	1	1	0	0	0	0	1	0
4 <sub>1</sub>	1	1	1	1	0	0	0	0	0
4 <sub>2</sub>	1	1	1	0	1	0	0	0	0
5	1	1	1	1	1	0	0	0	0
6	1	1	1	1	1	0	0	0	1
7	1	1	1	1	1	1	0	0	0

<sup>a</sup> 0 = primitive condition; 1 = derived.

port for the hypothesis that taxa E and F are more closely related to each other than to any other taxa, including taxon X. These two characters place limits on the possible positions of taxon X in any parsimonious interpretation of its relationships, but taxon X has missing entries for all characters that might further resolve its relationships within this region of the tree.

Parsimony analysis of the complete data including taxon X produces seven MPTs corresponding to the seven numbered positions in Figure 1a. In each of these MPTs, the relationships among the completely known taxa (A–F) are unaltered from those supported by analysis excluding taxon X. Thus, the MPTs differ only in the placement of taxon X relative to the constant relationships among the well-known taxa, with each placement corresponding to an equally parsimonious reconstruction of the missing entries in taxon X (Table 1) that requires no homoplasy.

Borrowing a term from philosophy, I call terminal taxa that display alternative equally parsimonious possible positions

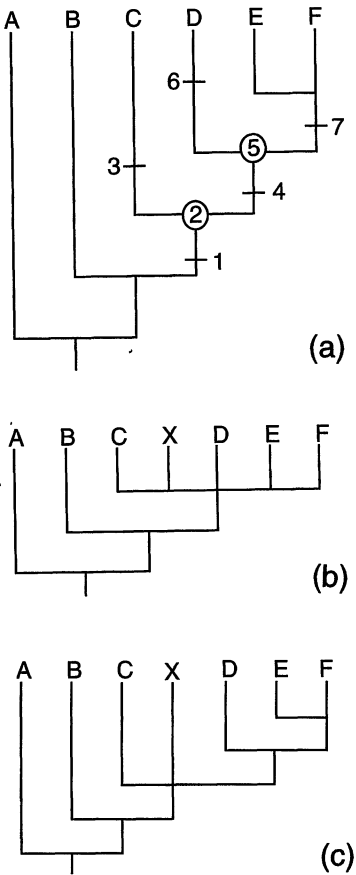


FIGURE 1. Hypothetical example based on the data in Table 1 showing the multiplicative effect of missing entries upon the number of most-parsimonious trees (MPTs) and the obfuscatory effect upon relationships indicated in consensus trees. (a) Unique MPT for the well-known taxa (A–F) showing the seven possible equally parsimonious positions of taxon X with respect to the unambiguous relationships among the other taxa. Each position corresponds to a separate MPT for the complete data. This tree is also a reduced Adams consensus tree for the full data. (b) Strict component consensus tree for the seven equally parsimonious trees. (c) Adams consensus tree for the seven equally parsimonious trees.

relative to the unambiguous relationships of other terminal taxa *underdetermined*. It is widely recognized that underdetermination may result from alternative equally parsimonious interpretations of homoplastic data. The hypothetical example illustrates the potential for underdetermination when there are missing entries but no ho-

moplasy (see also Wilkinson, 1995a). These two causes of underdetermination can be distinguished as positive (produced by contradictory evidence) or negative (produced by an absence of evidence).

As illustrated by the hypothetical example, the most salient practical consequence of including negatively underdetermined terminal taxa in parsimony analyses is an increase in the number of MPTs. Such increase may be dramatic, yielding many hundreds or thousands of MPTs. Most computer installations of parsimony analysis programs are limited in the number of trees that can be held in memory, and the number of MPTs for data sets with many taxa and an abundance of missing entries may exceed these limits.

Multiple MPTs present a problem for systematists when there is no additional basis for choosing among them. Consequently, consensus methods are widely used to summarize multiple MPTs. Given a set of fundamental trees, a consensus tree is intended to represent relationships that are common to all or to a majority of the fundamental trees. Although many consensus techniques have been developed, only a few are included in widely used parsimony analysis software, and each has particular limitations and drawbacks.

The limitations of the most popular consensus techniques can be illustrated with the hypothetical example of the seven MPTs (the fundamental trees) including the negatively underdetermined taxon X (Fig. 1a). Strict component consensus trees often suffer from a lack of resolution (Adams, 1986; Swofford, 1991; Wilkinson, 1992a, 1994). Thus, the strict component consensus tree (Fig. 1b) fails to represent the hypothesis that taxon D is more closely related to taxa E and F than to taxa A, B, and C, which is a common feature of the seven MPTs. Adams consensus trees (the second method of Adams, 1972) may be more fully resolved than strict component consensus trees, which reflects their much greater sensitivity to common structure (Adams, 1986; Wilkinson, 1994). Unfortunately, the polytomies of Adams consensus

trees have an ambiguous cladistic interpretation (Swofford, 1991; Wilkinson, 1992a, 1994). For example, in the Adams consensus tree (Fig. 1c), taxa C and X arise from the same polytomy and their possible relationships are indistinguishable. Thus although all the fundamental trees agree that taxa D, E, and F are more closely related to each other than they are to taxon C, this conclusion cannot be legitimately drawn from the Adams consensus tree alone. The same Adams consensus tree may result from cases where this conclusion is not true, for example if taxon C rather than taxon X or both taxon C and taxon X were negatively underdetermined.

These limitations of Adams and strict component consensus trees are related to a general difficulty for consensus trees that include underdetermined taxa. Attachment of an underdetermined taxon at any point on a consensus tree will either misrepresent the relationships of the underdetermined taxon or lead to a loss of information, either through loss of resolution or through ambiguity (Wilkinson, 1994). This problem is equally true of consensus trees based upon a majority rule and of those in which some relationship must be true of all fundamental trees to be included in the consensus. To summarize, negatively underdetermined taxa, which are often characterized by abundant missing entries, may lead to a proliferation of MPTs and necessitate the use of consensus methods, but commonly used consensus methods may fail to produce comprehensive summaries of relationships that are true of all the fundamental trees.

#### POTENTIAL SOLUTIONS

Just as the problem has two parts, so there are also two potential solutions. First, we could develop alternative consensus methods that are not subject to the general difficulties of those presently in widespread use. Such a solution would apply equally well to the limitations of these consensus methods in cases where numerous MPTs result from positive rather than negative underdetermination, homoplasy, or missing entries. A recently developed ap-

proach to this general consensus problem led to a family of reduced consensus methods in which underdetermined taxa may be excluded from reduced consensus trees facilitating the representation of relationships among taxa that are not underdetermined (Wilkinson, 1994). Reduced consensus trees can also be used as a framework for additional descriptions of the possible positions of excluded underdetermined taxa (Wilkinson, 1992a, 1994; Wilkinson and Benton, 1995). Figure 1a is a simple example of a reduced consensus tree that gives a better representation of relationships that are common to the seven MPTs than do the less resolved or ambiguous strict component and Adams consensus trees. This reduced consensus tree also provides a framework for the description of the possible positions of taxon X (indicated by numbered positions on the tree); it is also the single MPT for an analysis that does not include taxon X.

Algorithms for reduced consensus trees have been described (Wilkinson, 1994, 1995b) and used to resolve relationships among rhynchosaurs that are obscured by underdetermined taxa (Wilkinson and Benton, 1995). Programs that help implement this approach (Wilkinson, 1995c) are available from the author, but the methods are not included in any commonly used phylogenetic inference software. However, because data sets with abundant missing entries may support more MPTs than can be stored in memory, they may be particularly prone to the practical problem of sampling error to which no consensus method will be immune (i.e., consensus trees will be based on samples of MPTs rather than on all MPTs).

A second potential solution, *taxonomic reduction*, is more specific to the problems caused by missing entries and has been explored by several workers, including Benton (1990) and Nixon and Wheeler (1992). Identifying and removing negatively underdetermined terminal taxa from an analysis may enhance the resolution of consensus trees and thereby clarify relationships that are obscured when these underdetermined taxa are included in the analysis. It

is widely appreciated that adding and deleting taxa from an analysis may alter inferred relationships among the other taxa (e.g., Arnold, 1981; Donoghue et al., 1989), although as the hypothetical example illustrates this need not be the case. Consequently, any attempt to reduce the number of MPTs by deleting underdetermined taxa runs the risk of also changing the inferred relationships from those that are supported by the complete data. Ideally, we would only want to delete taxa if we were sure that the addition or deletion of the taxa would have no effect upon the inferred relationships of the other taxa included in the analysis. The aim is to produce more highly resolved inferences of these relationships, not to alter them.

Here, I develop a series of simple safe deletion rules that allow such inconsequential taxa to be identified and eliminated from an analysis in the secure knowledge that their elimination will not affect the inferred relationships of the remaining taxa. The effect of such taxa is simply to increase the number of MPTs. The reduction in the number of MPTs that accompanies the elimination of such taxa may increase the resolution and informativeness of strict component and Adams consensus trees for the reduced number of taxa and may also facilitate the construction of reduced consensus trees that are not subject to problems of incomplete sampling.

#### SAFE TAXONOMIC REDUCTION

##### *Taxonomic Equivalence*

Although strategies of taxonomic reduction have been used by previous workers, insufficient attention has been paid to the danger posed by the possibility that deleting taxa may alter inferred relationships among the remaining taxa. In contrast, the methods of safe taxonomic reduction are safe in the sense that only taxa that can have no effect upon the inferred relationships of other taxa included in the analysis are excluded. Consequently, safe taxonomic reduction should also have no effect upon tree lengths or consistency indices.

I call a group of terminal taxa that are

not demonstrably different with respect to phylogenetically informative characters *taxonomic equivalents*. Such taxa are said to display taxonomic equivalence. Taxonomic equivalence may be actual, in which case there are no missing entries and the equivalents share the same character states for all phylogenetically informative characters. Alternatively, the equivalence may be potential, when some comparisons cannot be made because there are missing entries for one or more of the equivalents.

Where equivalence is actual, the relationship between equivalents is necessarily symmetric. Where the equivalence is potential it may also be symmetric, i.e., when the equivalents have the same characters coded as missing and the same character states for all phylogenetically informative characters that are not coded as missing. However, for any pair of potentially equivalent taxa, there may be asymmetry, such that the missing data are concentrated in one of the pair; here the asymmetry is said to be all one way. This relation also pertains when there are missing entries for both taxa but only one of the pair has some character states that are known and scored but are coded as missing in the other taxon. Finally, where both taxa have missing entries for one or more characters that are scored in the other, then the asymmetry is said to be both ways. The various categories of equivalence are illustrated by the hypothetical taxa and character data in Table 2.

Previous strategies of taxonomic reduction have used the relative proportions of missing entries as a guide to choosing which taxa to exclude and which to include (e.g., Rowe, 1988; Benton, 1990). However, the absolute amount of missing data does not provide a faithful guide to the effects of addition and deletion of taxa upon relationships among other taxa. A better guide to whether taxa can be safely eliminated a priori is provided by the relations of taxonomic equivalence among the terminal taxa, which is affected both by the amount of missing data and, more importantly, by the relative distribution of these data.

TABLE 2. Hypothetical character data illustrating relations of taxonomic equivalence described in the text. Pairs of taxa demonstrate the various relationships. A and B: actual and therefore symmetric equivalence; all pairs except A and B: potential equivalence; C and D: symmetric potential equivalence; all except A and B, and C and D: asymmetric potential equivalence; D and E: asymmetric both ways; E and F, and A and C: asymmetric all one way.

Taxon	Characters					
	a	b	c	d	e	f
A	0	0	0	1	1	1
B	0	0	0	1	1	1
C	?	?	0	1	1	1
D	?	?	0	1	1	1
E	0	0	0	1	?	?
F	0	0	0	1	?	1

### Safe Deletion Rules

The development of safe deletion rules requires an understanding of the potential for the inclusion of taxa in an analysis to affect the inferred relationships among the other taxa. If a taxon can have no effect, then it can be safely deleted from the analysis. Studies of the effects of including fossil taxa in parsimony analyses have shown that these taxa can affect relationships when they bring particular new combinations of character states to an analysis (Donoghue et al., 1989). A minimum requirement for the inclusion of any terminal taxon to alter relationships among the other terminal taxa is that it must have unique combinations of phylogenetically informative characters. Only when this requirement is met can the addition of a taxon have any potential effect upon tree topology. The notion of taxonomic equivalence is significant here. Those taxa that are not equivalents must have unique combinations of characters. Thus, only those taxa that show taxonomic equivalence are potential candidates for elimination prior to analysis. Based upon the relations of taxonomic equivalence, a number of rather simple safe deletion rules can be developed.

The most obvious safe deletion rule applies when there is actual taxonomic equivalence. Here, the taxonomic equivalents each have the same combination of

character states. Such taxa will all originate from a common node in any MPT that does not include arbitrary resolutions. Thus, nothing can be gained from the inclusion of all the equivalents that cannot be gained from the inclusion of a single representative of them and the observation of actual taxonomic equivalence. Thus, the first safe deletion rule (rule 1a) is *represent all sets of actual taxonomic equivalents with a single taxon*. Which terminal taxon is retained is arbitrary because the equivalence is entirely symmetric. Although this is an obvious rule that has probably been applied by many systematists, there are many examples of published analyses that include actual taxonomic equivalents as separate terminal taxa (e.g., Tolson, 1987; Prum, 1988). Deleting actual taxonomic equivalents will not assist in ameliorating the problems of missing entries, but by reducing the number of terminal taxa, more MPTs may be sampled or exact rather than heuristic parsimony algorithms may be used.

When there is potential taxonomic equivalence, an analogous safe deletion rule applies, provided that the equivalence is also symmetric. Here again all the taxonomic equivalents have the same combinations of character states so that they all bring the same information to the analysis. The second rule (rule 1b) is *represent all sets of symmetric potentially equivalent taxonomic equivalents with a single taxon*. If sets of symmetric potential taxonomic equivalents are included in a parsimony analysis, they will not necessarily all originate from a single node on any MPT. Such an origin will be among the most-parsimonious alternatives, but the presence of missing data means that there may be equally parsimonious alternative placements of the underdetermined taxa that correspond to alternative equally parsimonious interpretations of their missing entries. Thus, rule 1b has greater potential for reducing the number of MPTs than does rule 1a. Rules 1a and 1b can be subsumed under the more general rule (rule 1): *represent all sets of symmetric taxonomic equivalents with a single taxon*.

Safe deletion rules can also be developed when there is asymmetric potential equivalence. The simplest case is when one of the set of potential taxonomic equivalents has no missing entries. In this case, none of the other potential equivalents can have character combinations that are not represented in the analysis by the taxon with no missing entries. A further safe deletion rule (rule 2a) is *remove any taxa that are potential taxonomic equivalents of taxa that have no missing entries*. If such potential taxonomic equivalents are included in an analysis, then their origin from the same node as their equivalent with no missing entries will be among the most-parsimonious hypotheses because there can be no more parsimonious reconstruction of the missing entries than that under which the taxa are identical to their potential equivalent with no missing entries. Under this interpretation, these taxa can add no more length to the tree than is already required by the inclusion of the taxon with no missing entries. However, there may be other equally parsimonious interpretations of the missing data and corresponding different positions of the underdetermined taxa. Thus, this rule has a greater potential for reducing numbers of MPTs than the rules based upon symmetry.

The safety of rule 2a is due to the fact that the asymmetry between the potential equivalents is all one way. Thus, the rule can be extended to cases when all the potential equivalents have missing entries but the asymmetry is all one way. The fourth safe deletion rule (rule 2b) is *when a pair of taxa are potential taxonomic equivalents and one taxon has no characters that are scored for it and that are coded as missing for the other taxa then remove the former*. Rules 2a and 2b can be generalized to rule 2: *eliminate those potential equivalents that have the greatest number of missing entries when the asymmetry is all one way*. The various deletion rules can be conceived of as different expressions of a general rule that identifies the potential for safe taxonomic reduction whenever there is taxonomic equivalence that is not asymmetric both ways.

#### SAURISCHIAN INTERRELATIONSHIPS

Gauthier (1986) used a data matrix of 84 polarized binary characters to infer phylogenetic relationships among 17 selected dinosaurian taxa, addressing in particular the relationships of Avialae (birds) within the Theropoda. Gauthier's data matrix is reproduced in Table 3. His parsimony analysis was performed in two parts using a mainframe (Michigan Terminal System [MTS]) installation of PAUP 2.4 (Swofford, 1985). First, he analyzed relationships among seven comparatively well-known taxa, producing a single MPT (Fig. 2a; Gauthier, 1986: fig. 8) with a length (L) of 94 steps and a consistency index (CI) of 0.89. The second analysis included the less well-known taxa and yielded numerous MPTs. Gauthier did not report tree statistics or numbers of MPTs for the second analysis, but he noted that (1986:8) "the sister group relationships among the seven well-known taxa were consistent across all possible trees."

Gauthier recognized that the multiple MPTs produced in his second analysis were a product of the abundant missing entries of the less well-known theropods. In addition, he claimed that the various resolutions of the relationships of the less well-known theropods were not supported by any evidence, implying that they represented only arbitrary resolutions (groups united by zero-length branches) produced by PAUP 2.4. Gauthier summarized the results of the second analysis with a "consensus" tree (of unspecified type) "incorporating polychotomies stemming from the levels supported by observable characters" (1986:8, fig. 9) (Fig. 2b). Despite Gauthier's description of the second analysis as including all taxa, one of the terminal taxa (*Hulsanpes*) is not included in this consensus cladogram, and its absence was not explained.

#### Reanalysis

Reanalysis of Gauthier's (1986) data provides a good illustration of the problems that may be caused by an abundance of missing entries and of the utility of safe



TABLE 3. Ingroup character data for the Saurischia (from Gauthier, 1986).

Terminal taxon	Abbrev. <sup>a</sup>	Characters <sup>b</sup>														
		1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-84						
Omnithischia	Orn	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0001
Sauropodomorpha	Sau	1110000000	0000000000	0000000100	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000001111	1110
Ceratosauria	Cer	1111111111	1111111111	1111111000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000001111	1111
Carnosauria	Car	1111111111	1111111111	1111111111	1111111111	1111000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000000000	0000001111	1?11
Ornithomimidae	Orn	0111111111	111?1111?1	1111111?1?	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	0000000111	1?11
Deinonychosauria	Dei	1111111111	1111111111	1111011111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1?11
Avialae	Avi	1001111?11	1111111111	0111011111	1111111101	1011??1111	1111111011	1111110111	1111110111	1111111111	1111111111	1111111111	1111111111	1111111111	1101111111	1?11
Ornitholestes	Ors	1?1?11?1??	?1?1?11111	11110?1111	?1?1?11111	?1?1?11111	?1?1?11111	?1?1?11111	?1?1?11111	?1?1?11111	?1?1?11111	?1?1?11111	?1?1?11111	?1?1?11111	?1?1?11111	111?
Coelurus	Coe	??????????	1?1?1?1?1?	?1?1?1?1?1?	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	????
Compsognathus	Com	11?111?1??	111?1111?1	?111011111	?111?1?111	?111?1?111	?111?1?111	?111?1?111	?111?1?111	?111?1?111	?111?1?111	?111?1?111	?111?1?111	?111?1?111	?111?1?111	????
Microenator	Mic	??????????	1?1?1?1?1?	?11?1?1?1?	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	?????????1	????
Sauromitholestes	Sas	?1?1?1?1??	1?1?11111?	?????1?1??	?????1?1??	?????1?1??	?????1?1??	?????1?1??	?????1?1??	?????1?1??	?????1?1??	?????1?1??	?????1?1??	?????1?1??	?????1?1??	1?1??
Hulsampes	Hul	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	??????????	????
Caenagnathidae	Cae	11111111??	?1?1111111	1111011?1?	?1?1111111	?1?1111111	?1?1111111	?1?1111111	?1?1111111	?1?1111111	?1?1111111	?1?1111111	?1?1111111	?1?1111111	011?1?1?1?	1?11
Elmsauridae	Elm	??????????	??????????	11?11111??	?1?1?1111?	11?11111??	?1?1?1111?	11?11111??	?1?1?1111?	11?11111??	?1?1?1111?	11?11111??	?1?1?1111?	?1?1?1111?	?????1?1?1	1?11
Procompsognathus	Pro	?1?1?1?1??	??????????	?111111?00	?????????0?	?????????0?	?????????0?	?????????0?	?????????0?	?????????0?	?????????0?	?????????0?	?????????0?	?????????0?	?????00?1?	????
Liliensternus	Lil	?1?1?1?1??	?1?11?1?11	1-111?110??	0?00?1?1?0	0000?1?1?0?	?????1?1?0?	?????1?1?0?	?????1?1?0?	?????1?1?0?	?????1?1?0?	?????1?1?0?	?????1?1?0?	?????1?1?0?	00000?1?1?0	????

<sup>a</sup> Abbreviations used in Figures 2-5.  
<sup>b</sup> 0 = primitive condition; 1 = derived.

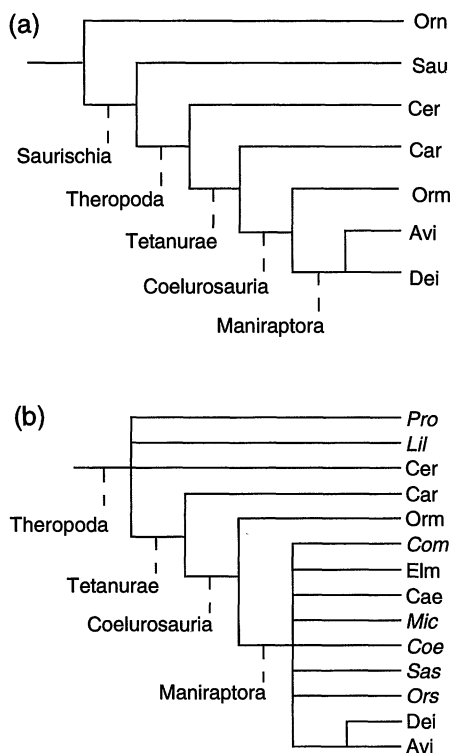


FIGURE 2. Cladograms showing saurischian interrelationships and classification as inferred by Gauthier (1986). Abbreviations as in Table 3. (a) First analysis. (b) Second analysis.

taxonomic reduction. The ingroup data (Table 3) contain over 41% missing entries, and for some of the individual terminal taxa, over 90% of the characters are coded as missing. In an initial reanalysis, MPTs were constructed for the complete data set using PAUP 3.1.1 (Swofford, 1993) and Hennig86 (Farris, 1988). Using Hennig86, 3,264 MPTs ( $L = 98$ ,  $CI = 0.86$ ) were found before there was no more memory available, and a strict component consensus tree (Fig. 3a) was produced using the only consensus method implemented in this program. Using PAUP 3.1.1, 3,700 MPTs ( $L = 98$ ,  $CI = 0.86$ ) were found before the memory was exhausted. The strict component consensus tree is the same as that produced in the Hennig86 analysis. My installation of PAUP 3.1.1 had insufficient memory to construct an Adams consensus tree for all 3,700 fundamental trees but

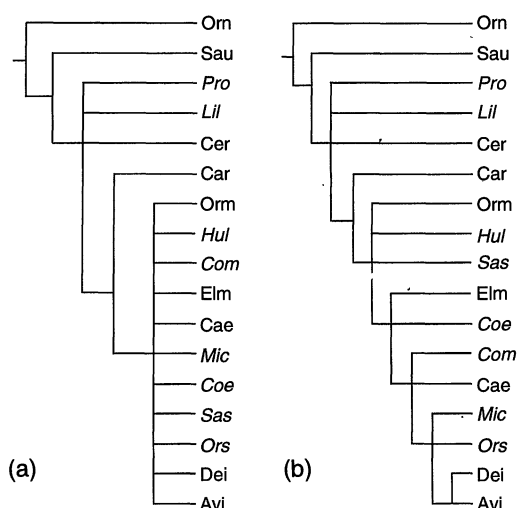


FIGURE 3. Reanalysis of Gauthier's (1986) saurischian data (Table 3) with PAUP 3.1.1. Abbreviations as in Table 3. (a) Strict component consensus tree for 3,700 most-parsimonious trees. (b) Adams consensus tree for 1,000 most-parsimonious trees.

was able to produce Adams consensus trees for nonoverlapping subsets of up to 1,000 fundamental trees, all of which produced the same Adams consensus tree (Fig. 3b).

The strict component and Adams consensus trees differ from Gauthier's consensus tree (Fig. 2b) in being less and more resolved, respectively. The most likely explanation for this is that the set of fundamental trees used as a basis for Gauthier's consensus tree was different from the set generated in the reanalysis. Gauthier did not report the number of MPTs found in his analysis, but the upper limit for the MTS installation of PAUP that he used was 300 (Nussbaum and Wilkinson, 1989). If Gauthier's consensus tree is a strict component consensus, then its greater resolution, compared with the strict component consensus based upon the over 3,000 MPTs found in the Hennig86 and PAUP 3.1.1 analyses, may be attributable to its being based upon a smaller sample of fundamental trees, i.e., the greater resolution may simply be a result of sampling error rather than a reflection of relationships

that are unambiguously supported by the parsimonious interpretation of the data.

Tree length, as determined by the reanalysis, is four steps longer than when the poorly known theropods are not included (Gauthier's first analysis). Thus, some additional character evolution must be posited when these taxa are included. Thus, any relationships among the poorly known theropods cannot be attributed entirely to zero-branch length arbitrary resolutions. The lower consistency index also demonstrates that the addition of the poorly known taxa increases character incongruence and implies additional homoplastic character evolution that could affect the inferred relationships among the well-known taxa. Similarly, the Adams consensus tree (Fig. 3b) suggests that there is considerably more agreement among the MPTs concerning resolution of the relationships of the less well-known theropods than is represented in either the strict component consensus or Gauthier's consensus. However, as noted earlier, there are problems in the interpretation of polytomies on Adams consensus trees. For example, from the Adams consensus tree it is not clear whether the *Elmsauridae* or *Coelurus* or both are taxa that are underdetermined and vary in position in the fundamental trees.

#### *Application of Safe Deletion Rules*

Application of rule 1a is not expected to ameliorate the problems caused by abundant missing entries because it applies only to taxa that have no missing entries. Furthermore, Gauthier's (1986) data set does not include any actual taxonomic equivalents. Despite its greater potential for reducing the number of MPTs, there are also no taxa in Gauthier's data set that satisfy the criteria for deletion employed in rule 1b. All taxonomic equivalence in the data is asymmetric.

Three terminal taxa, *Ceratosauria*, *Procompsognathus*, and *Liliensternus*, are a set of taxonomic equivalents in which *Ceratosauria* has no missing entries, *Procompsognathus* has 64 (76%) missing entries, and *Liliensternus* has 48 (57%) missing entries.

Both *Procompsognathus* and *Liliensternus* also show potential equivalence with other taxa. Applying rule 2a, these two terminal taxa were deleted and the reduced data set was analyzed using Hennig86 and PAUP 3.1.1, yielding 3,628 and 4,000 MPTs, respectively, before memory was exhausted. The MPTs do not differ in length (98) or consistency index (0.86) from those produced from the complete data set, as is expected if the removal of these taxa has no quantitative effect upon inferred character evolution and no qualitative effect upon relationships among the remaining taxa. Clearly, however, the limited selective deletions, applying rule 2a, do not achieve the aim of reducing the number of MPTs. In fact, the number of MPTs increased because the reduced number of taxa meant that each MPT occupied less computer memory than in the full analysis and hence more trees could be retained in the limited memory available.

There is one set of taxonomic equivalents, the *Deinonychosauria*, *Saurornitholestes*, and *Hulsanpes*, that allows the application of rule 2b. The last two taxa have no characters scored that are not also scored for the *Deinonychosauria*. The deletion of these taxa facilitated the use of an exact (branch and bound) parsimony algorithm. Using both Hennig86 and PAUP 3.1.1, 197 MPTs ( $L = 98$ ,  $CI = 0.86$ ) were found. This dramatic reduction in the number of MPTs circumvents the potential problems of sampling error. As expected, the safe deletion of these taxa had no effect on tree length or the consistency index.

A strict component consensus tree derived from the 197 fundamental trees (Fig. 4a) has two more nodes than the strict component consensus tree derived from the 3,700 MPTs for the unreduced data set (Fig. 3a). The increased resolution demonstrates the obfuscating effect of the inclusion of the deleted taxa, especially *Hulsanpes* and *Saurornitholestes*, when the strict component consensus method is used. The Adams consensus tree (Fig. 4b) differs from that derived from analyses of the complete data set (Fig. 3b) only in the removal of underdetermined taxa from poly-

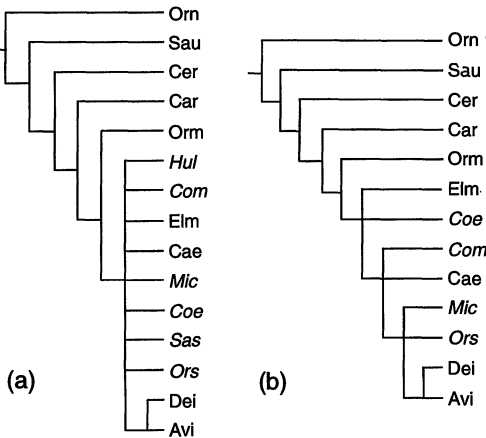


FIGURE 4. Reanalysis of Gauthier's (1986) saurischian data (Table 3) with PAUP 3.1.1 after deletion of *Procompsognathus*, *Liliensternus*, *Hulsanpes*, and *Saurornitholestes* through the application of safe deletion rules. Abbreviations as in Table 3. (a) Strict component consensus tree for 197 most-parsimonious trees. (b) Adams consensus tree for 197 most-parsimonious trees.

tomies. As a result, these ambiguous polytomies are transformed into unambiguous dichotomies, allowing greater resolution of the relationships that are obscured when the underdetermined taxa are included in the analyses.

The extent of taxonomic equivalence in the original data is shown as lists of all sets of taxonomic equivalents in Table 4. Each list begins with an index taxon, which shows some taxonomic equivalence, and the number of missing entries associated with that taxon. Subsequent entries include all taxa with which the index taxon shows taxonomic equivalence, and, for each of these taxa, the number of characters scored for that taxon but represented as missing in the index taxon. If this latter number is zero, there is scope for safe taxonomic reduction. Investigation of taxonomic equivalence is implemented in the TAXEQ program (Wilkinson, 1992b).

With the 13 ingroup terminal taxa remaining in the reduced data set, there remains much potential taxonomic equivalence, which is summarized in Table 5 and is simply a revised listing of the equiva-

lence shown in Table 4 for the remaining 13 taxa. For all sets of taxonomic equivalents, each member of the pair is scored for one or more characters that are scored as missing in the other taxon, i.e., there is always asymmetry both ways. Under these conditions, deletion of any terminal taxon from the data set cannot be guaranteed to be safe (i.e., to have no effect upon inferred relationships among the other taxa). Each of the remaining terminal taxa brings unique combinations of scored characters to the analysis that may affect cladogram topology. Also, some of the taxa with very high proportions of missing entries were retained in the analysis in preference to taxa with less missing data because the distribution and abundance of missing entries are taken into account in safe taxonomic reduction.

Further resolution can be obtained using a reduced consensus method (Wilkinson, 1994) to summarize relationships among the 197 MPTs from this final analysis and using this summary as a framework for the description of the possible positions of those taxa not included in the consensus (Fig. 5). There are several methods for determining possible positions of excluded taxa (Wilkinson and Benton, 1995), but they are beyond the scope of this paper and will not be discussed here.

## DISCUSSION

It has been recognized previously that an abundance of missing entries may bring special problems for phylogenetic inference using parsimony. Gauthier (1986) pointed out that in his analysis numerous MPTs were associated with the inclusion of poorly known taxa and thus with missing data, and this problem led him to adopt a strategy of taxonomic reduction. The strategy of taxonomic reduction employed by most workers faced with the problems of missing entries has been to eliminate those taxa that have more than a particular (and arbitrary) proportion of missing entries. This strategy was employed by Gauthier (1986), Benton (1990), Greenwald (1989), Novacek (1992a, 1992b),

TABLE 4. Table of taxonomic equivalence based upon the character data for Saurischia in Table 3. For each potential taxonomic equivalent of an index taxon, the figure in parenthesis is the number of characters scored for that taxon that are coded as missing in the index taxon.

Index <sup>a</sup>	MD <sup>b</sup>	Equivalents
Ceratosauria	0	<i>Procompsognathus</i> (0); <i>Liliensternus</i> (0)
Ornithomimidae	8	<i>Saurornitholestes</i> (1); <i>Hulsanpes</i> (0)
Deinonychosauria	6	<i>Saurornitholestes</i> (0); <i>Hulsanpes</i> (0)
<i>Ornitholestes</i>	40	<i>Coelurus</i> (6); <i>Microvenator</i> (10); <i>Saurornitholestes</i> (4); <i>Hulsanpes</i> (1); Caenagnathidae (15)
<i>Coelurus</i>	72	<i>Ornitholestes</i> (38); <i>Compsognathus</i> (37); <i>Microvenator</i> (9); <i>Saurornitholestes</i> (9); <i>Hulsanpes</i> (3); <i>Elmisauridae</i> (27); <i>Procompsognathus</i> (17)
<i>Compsognathus</i>	38	<i>Coelurus</i> (3); <i>Saurornitholestes</i> (4); <i>Hulsanpes</i> (1)
<i>Microvenator</i>	67	<i>Ornitholestes</i> (37); <i>Coelurus</i> (4); <i>Saurornitholestes</i> (9); <i>Hulsanpes</i> (2); <i>Elmisauridae</i> (26)
<i>Saurornitholestes</i>	72	Ornithomimidae (65); Deinonychosauria (66); <i>Ornitholestes</i> (36); <i>Coelurus</i> (9); <i>Compsognathus</i> (38); <i>Microvenator</i> (14); <i>Hulsanpes</i> (3); Caenagnathidae (40); <i>Elmisauridae</i> (22); <i>Procompsognathus</i> (18); <i>Liliensternus</i> (34)
<i>Hulsanpes</i>	81	Ornithomimidae (73); Deinonychosauria (75); <i>Ornitholestes</i> (42); <i>Coelurus</i> (12); <i>Compsognathus</i> (44); <i>Microvenator</i> (16); <i>Saurornitholestes</i> (12); Caenagnathidae (50); <i>Elmisauridae</i> (28); <i>Procompsognathus</i> (19)
Caenagnathidae	33	<i>Ornitholestes</i> (8); <i>Saurornitholestes</i> (1); <i>Hulsanpes</i> (2)
<i>Elmisauridae</i>	54	<i>Coelurus</i> (9); <i>Microvenator</i> (13)
<i>Procompsognathus</i>	64	Ceratosauria (64); <i>Coelurus</i> (9); <i>Saurornitholestes</i> (10); <i>Hulsanpes</i> (2); <i>Liliensternus</i> (26)
<i>Liliensternus</i>	48	Ceratosauria (48); <i>Saurornitholestes</i> (10); <i>Procompsognathus</i> (10)

<sup>a</sup> Taxon that shows potential taxonomic equivalence.

<sup>b</sup> Number of missing entries in the index taxon.

and Rowe (1988). However, as Novacek (1992a:75) noted,

Although this prescription has some logic, it is often difficult to predict the effectiveness of taxa based on their amount of character information. . . . The kinds of characters preserved, not just the degree of character representation, account for the potential influence of an added taxon.

Thus, it is not simply the proportion of missing entries that directly determines whether the deletion of taxa will have any effect upon the inferred relationships

among the remaining taxa. Rather, the important factor is the uniqueness of the combinations of character states that are present in a taxon. Both Norell and de Queiroz (1991) and Wilkinson and Benton (1995) achieved reductions in numbers of MPTs by incorporating poorly known fossil taxa into their data sets. The strategy of safe taxonomic reduction developed here allows the identification of those taxa that can be deleted from an analysis without any possible effect upon the results of the

TABLE 5. Table of taxonomic equivalence for Gauthier's (1986) saurischian data in Table 3 after deletion of taxa that satisfy the conditions of the safe deletion rules. For each potential taxonomic equivalent of an index taxon, the figure in parenthesis is the number of characters scored for that taxon that are coded as missing in the index taxon.

Index <sup>a</sup>	MD <sup>b</sup>	Equivalents
<i>Ornitholestes</i>	40	<i>Coelurus</i> (6); <i>Microvenator</i> (10); Caenagnathidae (15)
<i>Coelurus</i>	72	<i>Ornitholestes</i> (38); <i>Compsognathus</i> (37); <i>Microvenator</i> (9); <i>Elmisauridae</i> (27)
<i>Compsognathus</i>	38	<i>Coelurus</i> (3)
<i>Microvenator</i>	67	<i>Ornitholestes</i> (37); <i>Coelurus</i> (4); <i>Elmisauridae</i> (26)
Caenagnathidae	33	<i>Ornitholestes</i> (8)
<i>Elmisauridae</i>	54	<i>Coelurus</i> (9); <i>Microvenator</i> (13)

<sup>a</sup> Taxon that shows potential taxonomic equivalence.

<sup>b</sup> Number of missing entries in the index taxon.

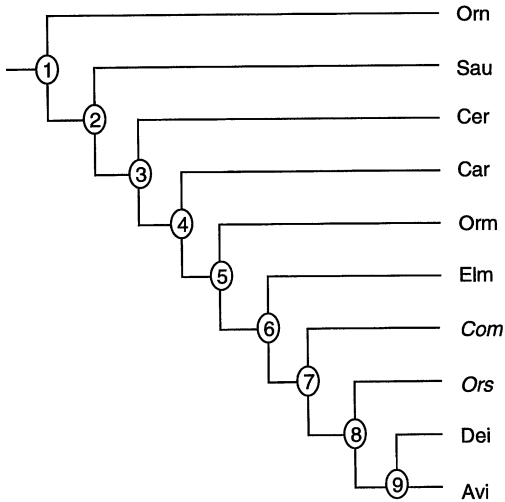


FIGURE 5. Reduced consensus tree for the Saurischia based on Gauthier's (1986) data (Table 3). The possible positions of excluded taxa are as follows. *Caenagnathidae*: between nodes 6 and 7, between nodes 7 and 8, above node 8, and below node 9; *Microvenator*: between nodes 6 and 7, above node 7, and below node 9; *Coelurus*: between nodes 5 and 6, above node 6, and below node 9; *Procompsognathus* and *Liliensternus*: between nodes 2 and 3, above node 3, and below node 4; *Hulsanpes* and *Saurornitholestes*: between nodes 4 and 5 and above node 5.

analyses and thereby removes the danger of deleting taxa that may be crucial to the parsimonious interpretation of relationships based upon the available data.

Crepet and Nixon (1989) investigated the phylogenetic relationships of some fossil Fagaceae by comparing their character states with those of the terminal taxa and nodes of MPTs that were based only on extant genera. These authors also argued that if the inclusion of a taxon adds no length to the cladogram then it cannot affect the relationships among the other taxa, and they used this assumption to evaluate the possible effects of the excluded fossil taxa. Similarly, Nixon and Wheeler (1992: 134) suggested that in some cases it is "better to exclude the fossil from the analysis and then determine the range of nodes (groups) to which the fossil can be attached without affecting the length of the cladogram."

An experimental approach to safe taxonomic reduction involving multiple anal-

yses to identify those taxa that may be eliminated without altering tree length might provide an alternative basis for safe taxonomic reduction. However, the central idea that having no effect upon tree length implies no effect upon tree structure does not always hold. Novacek (1992b) described a case in which the addition of taxa had no effect upon tree length but did affect tree topology. In Novacek's example, parsimony analysis of the relationships among four extant eutherian groups produced two MPTs reflecting alternative optimizations of homoplastic characters. In a second analysis, the inclusion of three fossil taxa added nothing to tree length but yielded a unique MPT, consistent with only one of the MPTs for the extant taxa alone. In this case, although the fossil taxa added no additional homoplasy, they did affect the optimization of the homoplastic characters and rendered one of the two initial tree topologies less parsimonious than the other. For this reason, determining that there is no difference in tree length when a taxon is included or excluded is not sufficient to prove that the elimination of a taxon is safe.

Whenever cladistic data sets include taxonomic equivalents that are not asymmetric both ways, there is scope for safe taxonomic reduction using safe deletion rules. As the example shows, this simple approach can greatly reduce numbers of MPTs when there is much missing data and thereby can facilitate the construction of more informative summaries of the relationships that are supported by the most-parsimonious interpretation of the complete data. In addition, safe taxonomic reduction may (1) reduce the duration of subsequent analyses, (2) allow exact rather than heuristic methods to be employed, and (3) reduce or overcome problems of incomplete sampling of MPTs.

#### ACKNOWLEDGMENTS

I am grateful to Nick Arnold, Mike Benton, Peter Forey, Rod Page, Glenn Storrs, Dave Unwin, Paul Upchurch, Matthew Wills, and two anonymous reviewers for discussions, reviews of manuscripts, and encouragement. The TAXEQ and REDCON 2.0 programs and documentation are available from the author

upon receipt of a blank disc (PC formatted). This work was supported in part by SERC grant GR/F 87912.

## REFERENCES

- ADAMS, E. N. 1972. Consensus techniques and the comparison of taxonomic trees. *Syst. Zool.* 21:390–397.
- ADAMS, E. N. 1986. N-trees as nestings: Complexity, similarity and consensus. *J. Classif.* 3:299–317.
- ARNOLD, E. N. 1981. Estimating phylogenies at low taxonomic levels. *Z. Zool. Syst. Evolutionsforsch.* 19: 1–35.
- BENTON, M. J. 1990. The species of *Rhynchosaurus*, a rhynchosaur (Reptilia, Diapsida) from the Middle Triassic of England. *Philos. Trans. R. Soc. Lond. B* 328:213–306.
- BRYANT, H. N., AND A. P. RUSSELL. 1992. The role of phylogenetic analysis in the inference of unpreserved attributes of extinct taxa. *Philos. Trans. R. Soc. Lond. B* 337:405–418.
- CRÉPET, W. L., AND K. C. NIXON. 1989. Earliest megafossil evidence of Fagaceae: phylogenetic and biogeographic implications. *Am. J. Bot.* 76:842–845.
- DONOGHUE, M. J., J. A. DOYLE, J. A. GAUTHIER, A. G. KLUGE, AND T. ROWE. 1989. The importance of fossils in phylogeny reconstruction. *Annu. Rev. Ecol. Syst.* 20:431–460.
- FARRIS, J. S. 1988. Hennig86, version 1.5. Distributed by the author, Port Jefferson Station, New York.
- GAUTHIER, J. A. 1986. Saurischian monophyly and the origin of birds. *Mem. Calif. Acad. Sci.* 8:1–47.
- GREENWALD, N. 1989. Effects of missing data and homoplasy on estimates of multituberculate phylogeny. *J. Vertebr. Paleontol.* 9:24A.
- NIXON, K. C., AND Q. D. WHEELER. 1992. Extinction and the origin of species. Pages 119–143 in *Extinction and phylogeny* (M. J. Novacek and Q. D. Wheeler, eds.). Columbia Univ. Press, New York.
- NORELL, M. A., AND K. DE QUEIROZ. 1991. The earliest iguanine lizard (Reptilia: Squamata) and its bearing on iguanine phylogeny. *Am. Mus. Novit.* 2997:1–16.
- NOVACEK, M. J. 1992a. Fossils as critical data for phylogeny. Pages 46–88 in *Extinction and phylogeny* (M. J. Novacek and Q. D. Wheeler, eds.). Columbia Univ. Press, New York.
- NOVACEK, M. J. 1992b. Fossils, topologies, missing data, and the higher level phylogeny of eutherian mammals. *Syst. Biol.* 41:58–73.
- NUSSBAUM, R. A., AND M. WILKINSON. 1989. On the classification and phylogeny of caecilians (Amphibia: Gymnophiona), a critical review. *Herpetol. Monogr.* 3:1–42.
- PLATNICK, N. L., C. E. GRISWOLD, AND J. A. CODDINGTON. 1991. On missing entries in cladistic analysis. *Cladistics* 7:337–343.
- PRUM, R. O. 1988. Phylogenetic interrelationships of the barbets (Aves: Capitonidae) and toucans (Aves: Ramphastidae) based on morphology with comparisons to DNA–DNA hybridization. *Zool. J. Linn. Soc.* 92:313–343.
- ROWE, T. 1988. Definition, diagnosis, and origin of Mammalia. *J. Vertebr. Paleontol.* 8:241–264.
- SWOFFORD, D. L. 1985. PAUP: Phylogenetic analysis using parsimony, version 2.4. Illinois Natural History Survey, Champaign.
- SWOFFORD, D. L. 1990. PAUP: Phylogenetic analysis using parsimony, version 3.0. Illinois Natural History Survey, Champaign.
- SWOFFORD, D. L. 1991. When are phylogeny estimates from molecular and morphological data incongruent? Pages 295–333 in *Phylogenetic analysis of DNA sequences* (M. M. Miyamoto and J. Cracraft, eds.). Oxford Univ. Press, New York.
- SWOFFORD, D. L. 1993. PAUP: Phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign.
- TOLSON, P. J. 1987. Phylogenetics of the boid snake genus *Epicrates* and Caribbean vicariance theory. *Occas. Pap. Mus. Zool. Univ. Mich.* 715:1–68.
- WILKINSON, M. 1992a. Consensus, compatibility and missing data in phylogenetic inference. Ph.D. Thesis, Univ. Bristol, Bristol, England.
- WILKINSON, M. 1992b. TAXEQ, software and documentation. Univ. Bristol, Bristol, England.
- WILKINSON, M. 1994. Common cladistic information and its consensus representation: Reduced Adams and cladistic consensus trees and profiles. *Syst. Biol.* 43:343–368.
- WILKINSON, M. 1995a. Arbitrary resolutions, missing entries, and the problem of zero-length branches in parsimony analysis. *Syst. Biol.* 44:108–111.
- WILKINSON, M. 1995b. More on reduced consensus methods. *Syst. Biol.* 44:435–439.
- WILKINSON, M. 1995c. REDCON 2.0: Reduced consensus programs and documentation. Univ. Bristol, Bristol, England.
- WILKINSON, M., AND M. J. BENTON. 1995. Missing data and rhynchosaur phylogeny. *Hist. Biol.* 10:137–150.

Received 28 February 1994; accepted 23 March 1995